Introduction

Mexico is a megadiverse country in terms of ecosystems, biological organisms, and cultures. In this context, the production of traditional agave spirits blends the variety of local biological resources (agave plants and microorganisms), with the heterogeneity of traditional autochthonous practices, to ferment and distill cooked agave hearts (Fig. 1C). Most of the artisanal mezcal production relies on open fermentations performed by yeasts and bacteria living between the factories and the variety of surrounding ecosystems.

To study the evolutionary relationship, genetic diversity, and population structure of yeasts employed in production of agave spirits, we isolated and sequenced S. cerevisiae strains from fermenting agave juice of 62 different mezcal factories. Then, we compiled a dataset of 361 S. cerevisiae genomes:

- 154 from our agave collection
- 23 from tequila reported in Lachance, 1995 and Padilla et al., 1994
- 168 reported in Peter et al., 2018, (whole Mexico and French Guiana)
- 16 from Brazil reported in Barbosa et al., 2016

Results

Most mezcal agave strains group in a new cluster

Fig. 2) A. Maximum likelihood tree using the program RAxML with the subset of 361 genomes using 961,937 SNPs. The tip branches with agave strains are colored based on the isolation region depicted in the inset. 2B. Multidimensional scaling analysis using the program plink v1.9. Graph of the 1st and 2nd components, using 29,680 biallelic SNPs from 1,017 S. cerevisiae genomes, showing that the agave strains form a separate group.

Highlights

- The genomes of most mezcal S. cerevisiae strains group in a new cluster.
- Some migrant strains isolated from Tequila, group with bread strains, and interestingly some others show signatures of admixture between bread and West mezcal strains.
- Population structure analyses show that genomic variation correlates with the geographical regions with some admixture between subpopulations.
- Introgressions of S. paradoxus in these Mexican Agave strains are prevalent. We even found one mezcal strain with more than 370 S. paradoxus genes.

Fig. 3) A. Inference of population structure using the programs Admixture and porg on 173 S. cerevisiae Mexican agave, bread and French Guiana isolates with a set of 36,714 markers filtered by LD with 500 bp and r2<0.3, assuming K=11 subpopulations, with 15 random seeds. Each rectangle on the x-axis represents one strain, and its colors represent the fraction of the genetic material in each strain assigned to each of the 11 clusters. 3B. Principal component analysis of the same strains and markers as panel 3A. 3C. Number of genes from S. paradoxus for each strain, separated by regions.

Perspectives

- Estimates of genetic diversity per group are work in progress.
- We also plan to evaluate the impact of introgressions in our findings.
- Increase the sample size of tequila isolates.

Acknowledgments

We thank all agave fermentation producers, master mezcalers, farmers and distillery workers for their hospitality and generosity.

References